

Shwedha Srinivasan

ssrinivasan2@hawk.iit.edu | +1 (773) 837-7430 | [LinkedIn](#) | [GitHub](#)

SUMMARY

Individual with skills in the field of Python, SQL, storytelling and data visualization and various other disciplines of data science and machine learning that equips to effectively communicate complex information, and translating insights for diverse audiences.

EDUCATION

Illinois Institute of Technology

Master of Data Science (Major GPA: 3.714, Overall GPA: 3.4)

Relevant coursework: Database Organisation, Machine Learning, Data Mining, Applied Statistics, Big Data, Data preparation and Analytics, Project Management, Monte Carlo Methods in Finance.

Madras Institute of Technology

Bachelors of Engineering, Electronics and Communication

August 2018 – May 2022

PROJECTS

Customer Segmentation and Churn Prediction

- Addressed feature selection and preprocessing to prepare the dataset for analysis. Utilized K-Means clustering algorithm and RFM analysis to segment customers into distinct groups based on their behaviour and preferences.
- Presented different models to predict churn rate where the best model was with an accuracy of 80.15%.
- Employed machine learning algorithms including Random Forest Classifier, Logistic Regression, GaussianNB, Decision Tree Classifier, and XGB Classifier.

Bitcoin Prediction

- Analysed extensive historical Bitcoin price data to identify crucial patterns and trends. Employed sentiment analysis to gauge market sentiments, enriching Time Series Forecasting accuracy.
- Developed a robust ARIMA model to predict Bitcoin prices for the upcoming 20 days, achieving 83% accuracy.
- Utilized diverse data science technologies, including ETL, statistical learning, Python (CNN, LSTM, ARIMA, Linear regression, Random forests), sentiment analysis, and time series forecasting to provide valuable insights and inform decision-making.

IPL live score prediction

- Predicted scores of both teams playing at the end of each powerplay in live i.e., when sports competition took place.
- Identified scores with the help of Random Forest classifier in python with accuracy 80.15%.
- Technologies used: EDA, Classifiers like Random Forest, lasso, ridge, XGBoost, etc.

WORK EXPERIENCE

Skilbi Corp

Data Science Intern

Chicago, IL

October 2023 – December 2023

- Administered textual datasets, and classified them into clusters using vector embedding with an increased accuracy of 80%.
- Improved the career recommendation paths with the help of OpenAI.

VOLUNTEER EXPERIENCE

Calamos Investments (Part of Data Science Practicum course)

Chicago, IL

- Collaborated with marketing teams to validate the impending product launch through insightful data analysis.
- Conducted Exploratory Data Analysis (EDA) on a vast dataset of 60 million records, extracting pivotal insights aimed at boosting monetization across diverse portfolio categories.
- Utilized clustering analysis and visualisation techniques to identify commonalities among various firms, facilitating strategic decision-making.

University of British Columbia

Undergraduate research mentee

Remote

- Engaged in a collaborative research project with a UBC mentor, focusing on training a machine learning model using publicly available data. The objective was to discern patterns elucidating gene regulation influenced by DNA methylation.
- Integrated with other student Mentees and a Mentor in the REX program to conduct literature-based research on patient's data who suffered from DNA methylation loss in colorectal cancer genes and generate a rudimentary cleaning of data from the intronic enhancers.

LEADERSHIP EXPERIENCE

Food Secretary

Residential Community at Madras Institute of Technology

India

May 2021 – June 2022

- Earned a responsibility for checking food ingredients utilized for making food for students and residential staffs.
- Evaluated bills for purchase of goods and recorded the same to pass to higher officials.
- Developed a sense of leadership, team work, communication, and presence of mind.

SKILLS & INTERESTS

Data analysis and Visualization: Python (pandas, NumPy, matplotlib, scikit-learn), Excel, R, EDA, Tableau.

Predictive Modelling: Regression, Classification, Clustering, ETL, Deep learning and Neural networks.

Mathematical Skills: Statistical Analysis (A/B testing), Probability Theory, Bayesian Theory, Gaussian Process, Stochastic Process.

Data Handling: MySQL, PostgreSQL, PhpMyAdmin, AWS.

Certifications: Microsoft (DAT256x: Essential Math for Machine Learning: Python Edition), 365 Data Science certification, ISRO (Machine learning to Deep Learning).

Others: Adobe Illustrator, Microsoft PowerPoint, XAMP, Android studio.